

# In-silico modelling of reaction networks involved in prebiotic chemistry to understand the origins of life

SIDDHANT SHARMA<sup>1</sup>, AAYUSH ARYA<sup>1,2</sup>, JESSICA RAY<sup>1</sup>, EDUARDO ALEJANDRO LOZANO GARCIA<sup>1,3</sup>, ROMULO LEONCIO CRUZ SIMBRON<sup>1,4</sup>, JAKOB LYKKE ANDERSEN<sup>5</sup>, RANA DOGAN<sup>1,6</sup>, HUAN CHEN<sup>7</sup> AND HENDERSON JIM CLEAVES<sup>1,8,9</sup>

<sup>1</sup>Blue Marble Space Institute of Science

<sup>2</sup>Lovely Professional University

<sup>3</sup>Instituto Politécnico Nacional

<sup>4</sup>Universidad Nacional de Ingeniería

<sup>5</sup>University of Southern Denmark

<sup>6</sup>Bogazici University

<sup>7</sup>National High Magnetic Field Laboratory

<sup>8</sup>Earth-Life Science Institute, Tokyo Institute of Technology

<sup>9</sup>Institute for Advanced Study

Presenting Author: siddhaantsharma.ss@gmail.com

Complex chemical reaction networks can grow exponentially in terms of the chemical diversity they generate. It is unknown whether such networks easily discover or shuttle fluxes through autocatalytic sub-networks. In general, such sub-networks may be common or rare or anywhere in between in organic chemistry. We aim to provide a map for experimental chemists studying complex organic reactions using an automated rule-based reaction generation to simulate the reactions involved in various plausible abiotic reactions proposed to account for the organic diversity observed in carbonaceous meteorites, thus providing enough data for ground-truthing of our methods. We applied graph transformation rules based on well-documented reaction mechanisms, chemical intuition and applied various constraints to the outputs, such as disallowed output structural motifs, thereby restricting them. We used isomorphism tests to match the output molecular structures to experimentally reported structures as a test of completeness of the methods in our study. The monoisotopic exact masses of the molecules in the computed reaction network product set were calculated and used to match peaks identified in high-resolution FT-ICR-MS data of the same reaction. We modeled the alkaline degradation of glucose using our workflow and found that our model was able to explain 96% of the structures reported in analytical studies (e.g., Yang and Montgomery, 1996). When the same workflow was applied to simulate formose chemistry, we were able to match all the structures reported by Decker & Schweer (1982) and Omran et al. (2020). The reaction network was further assessed for the existence of potentially autocatalytic loops by loading the network topology into a graph database where pattern matching queries could be executed to search for patterns of interest. This work demonstrates some efficient methods for finding reaction pathways and autocatalysis in *In silico* modeled reaction networks. This kind of *In silico* modeling enables the comprehensive study of chemical reaction pathways and, the

knowledge of possible compounds involved in such reaction networks can guide future untargeted searches for the analysis of organics in the cometary, meteorite, and extraterrestrial planetary samples and provide new insights into the mysteries of the origin of life.

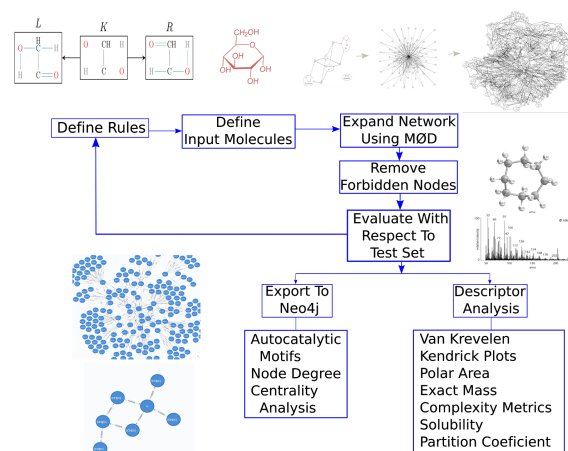


Figure 1. Systematic workflow for the methods described in our study.

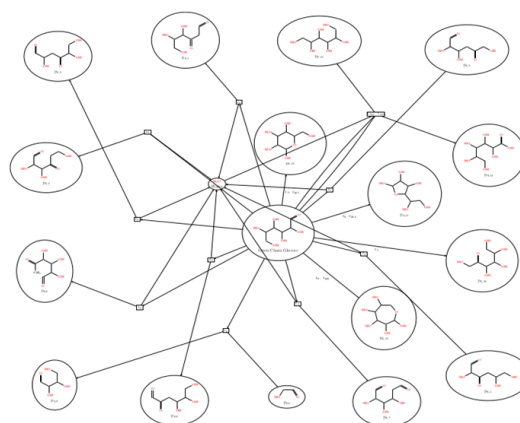


Figure 2. Sample glucose network generated by our methods.